

受 験 番 号					

問題冊子にも受験番号を書きなさい。

東京大学大学院新領域創成科学研究科

情報生命科学専攻

平成17(2005)年度大学院入学試験問題

修士・博士後期課程

専 門 科 目

平成15年8月25日(水)

10:30~12:30(120分)

注意事項

1. 試験開始の合図があるまで、この冊子を開いてはいけません。
2. 本冊子の総ページ数は12ページです。落丁、乱丁、印刷不鮮明な箇所などがあった場合には申し出ること。
3. 解答には必ず黒色鉛筆(または黒色シャープペンシル)を使用しなさい。
4. 問題は7題出題されます。問題1~3の3問と問題4~7から選択した2問の合計5問に解答しなさい。
5. 解答用紙は計5枚配られます。各問題に必ず1枚の解答用紙を使用しなさい。解答用紙に書ききれない場合は、裏面にわたってもよい。
6. 解答は日本語または英語で記入しなさい。
7. 解答用紙の指定された箇所に、受験番号と選択問題の場合はその用紙で解答する問題番号を記入しなさい。問題冊子にも受験番号を記入しなさい。
8. 草稿用紙は本冊子から切り離さないこと。
9. 解答に関係ない記号、符号などを記入した答案は無効とします。
10. 解答用紙・問題冊子は持ち帰ってはいけません。

(このページは草稿用紙として使用してよい)

問題1：下記の(1)～(4)の問題に答えよ。

(1) n 次正方行列 A が正則であるとは、 $XA = AX = I$ (I は単位行列) を満たす行列 X が存在することである。 X を逆行列といい、 A^{-1} とあらわす。 $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ とする。

問1 - 1 . A が正則である必要十分条件を示し、逆行列 A^{-1} を求めよ。

問1 - 2 . A が正則でない場合に固有値のひとつは0であることを示せ。

問1 - 3 . $a=5, b=4, c=-2, d=-1$ のとき A のすべての固有値とそれぞれに対応する固有ベクトルを求めよ。

(2) 16進数では10進数をそれぞれ以下のように表記するとして次の問いに答えよ。
 $A=10, B=11, C=12, D=13, E=14, F=15$

問2 - 1 . 2進数 1001001 は10進数ではいくらか。

問2 - 2 . 2進数 11111101 は8進数ではいくらか。

問2 - 3 . 16進数 D3F は2進数ではいくらか。

(3) 以下の設問に (論理和) (論理積) (全称) (存在) \neg (否定)、(含意) などを用いて答えよ。

問3 - 1 . 命題論理に関して以下の空欄 (アンダーライン) を埋めた式を書け。

(a) $F \ (G \ H) = (F \underline{\quad} G) \underline{\quad} (F \underline{\quad} H)$

(b) $\neg (F \ G) = \neg F \underline{\quad} \neg G$

問3 - 2 . 述語論理において、 $Q(x)$ と $R(x)$ をそれぞれ、「 x は有理数である」、「 x は実数である」を表す述語記号とする。そのとき、以下の3つを、 Q と R を含む式で書け。

(c) すべての有理数は実数である

(d) ある実数は有理数である

(e) すべての実数が有理数とは限らない

(4) テキストとパターンの2つの文字列が与えられたときに、テキスト上にパターンが存在するか否かを調べる文字列照合問題を考える。テキストとパターンの先頭文字から順に一致するか調べ、不一致の場合、パターンを一文字右にずらして同じことを繰り返すという単純な方法を用いた場合について、以下の問いに答えよ。

問4 - 1 . テキストの長さを N 、パターンの長さを M ($N > M$) としたときに、最悪 (もっとも比較回数が多くなる) の場合の文字の比較回数を N と M を使って表せ。

問4 - 2 . そのような最悪の場合が起こる、テキストとパターンの例を示せ。

問題 2 : 以下の問 1 ~ 3 に答えよ。

問 1 . 以下の問題文中 (1) から (1 0) に適当な語を入れよ。

- ・ DNA のクローニングには、特異的な塩基配列を認識して DNA の二本鎖を切断する (1) と DNA 断片をつなぎ合わせる (2) の 2 種類の酵素が重要な役割を果たす。
- ・ 微量の DNA から特異的な領域を増幅するためには、2 種類のプライマーと耐熱性 DNA 合成酵素を利用した (3) という方法が有効である。
- ・ mRNA からタンパク質への翻訳は、3 ヌクレオチドを単位とする (4) をアミノ酸に対応させることにより行われ、(5) をコードする AUG から翻訳が開始される。
- ・ タンパク質は、アミノ酸が (6) 結合により連結してできている。
- ・ タンパク質には、合成後に細胞膜や特定の細胞内小器官に移行するための (7) と呼ばれる領域をもつものがある。
- ・ 化学反応を触媒する能力をもつ RNA を (8) という。
- ・ 1 分子のグルコースを完全酸化することにより 3 0 分子の (9) が生成する。真核細胞では、この酸化反応は主に (1 0) で行われる。

問2 . 以下の左の A~J の各単語に対し、最も関連が深いものを右の単語群から1つずつ選んで番号を記せ。

- | | |
|----------|-----------|
| A グリア | 1 アクチン |
| B コルヒチン | 2 イオン |
| C セントロメア | 3 タウリン |
| D チャンネル | 4 カドヘリン |
| E リガンド | 5 チューブリン |
| F リソソーム | 6 ミエリン |
| G リボソーム | 7 ゴルジ体 |
| H 抗体 | 8 シトクローム |
| I 細胞接着 | 9 ヒストン |
| J 糖鎖修飾 | 10 プロモーター |
| | 11 レセプター |
| | 12 リンパ球 |
| | 13 小胞体 |
| | 14 染色体 |
| | 15 分解酵素 |

問3 . 以下の各項目についてそれぞれ2行以内で簡潔に説明せよ。

- 1 . 細胞膜の基本構造
- 2 . 岡崎フラグメント
- 3 . 神経シナプス
- 4 . 真核生物の細胞周期

問題3：以下の(1)～(5)に答えよ。

(1) 以下の語の組をそれぞれ100文字以内で解説しなさい。

1. オースログ遺伝子とパラログ遺伝子
2. SNPとマイクロサテライト

(2) 括弧の中に入る適切な語を答えなさい。

1. 20種類のアミノ酸残基の物理化学的な類似性を数値化する指標として(1)や(2)が使われる。
2. 相同な遺伝子のDNA塩基配列を、アミノ酸をコードする3つの塩基を単位として比較すると、変わりやすいのは(3)番目、(4)番目、(5)番目の塩基の順である。

(3) 以下の3つの数を答えなさい。

1. 5塩基の一本鎖DNAで、可能な配列の数
2. 5塩基対の二本鎖DNAで、可能な配列の数
3. 6塩基対の二本鎖DNAで、可能な配列の数

(4) アミノ酸配列データベースの中から、問い合わせアミノ酸配列と類似したエントリを見つける処理(配列相同性検索)を行うBLASTプログラムについて、1～3から正しいものを全て選べ。

1. 動的計画法(ダイナミック・プログラミング)を用いて、Smith-Watermanアルゴリズムを実装したものである。
2. 問い合わせアミノ酸配列を文字列にエンコードして扱い、アミノ酸間の特性や進化的な関係の影響を回避してアラインメント及び検索を行う。
3. アラインメントの統計的な有意さを期待値(E-value)で評価する。

(5) 塩基配列 ATGT と ACCTA のアラインメント結果

ACCT-A

A--TGT

を表現するマトリクスは、マッチした塩基対を |、ミスマッチを x、ギャップを - と | で示すと以下のようなになる。

	A	C	C	T	A
A		-	-		
T					
G					
T					x

以上を参考に、次のアラインメント結果を示すマトリクスのなかで、アラインメントを正しく表現していないものを全て選び、その理由を述べよ。

a.

-				
		x		
				x

b.

	-			
		x	x	

c.

	-	-	-	x

d.

x				
		-	-	
				x

問題 4 : すべて値が異なる N 個 ($N > 2$) の要素の列 $x(1), x(2), \dots, x(N)$ を大きい順 (降順) に並べ替えるソーティング問題について以下の問いに答えよ。

問 1 . $x(1), x(2), \dots, x(N)$ から最も大きな要素を見つけるのに、値の比較は何回必要か。

問 2 . $x(1), x(2), \dots, x(N)$ のうち最も大きな要素から最も小さな要素まで 1 個ずつ順に見つけるために、問 1 と同様の手続きを繰り返した場合、値の比較は何回必要か。

問 3 . $N=2^k$ のとき、 $x(1), x(2), \dots, x(N)$ を要素の個数が同じ 2 グループに分ける操作を再帰的に行うとする。それぞれのグループが 1 個の要素となるまで分割するとき、合計何回の分割操作が必要か。

問 4 . $x(1), x(2), \dots, x(N)$ が m 個と n 個の要素からなる 2 グループに分けられ ($N = m + n$)、それぞれのグループの要素は降順にソーティングされているとする。この 2 つのグループが既に降順にソーティングされていることを利用して、2 つのグループの要素を順に比較すると、最悪の場合でも ($N - 1 = m + n - 1$) 回の値の比較によって、($N = m + n$) 個の要素をソーティングできる。この方法と問 3 の分割を利用して ($N=2^k$) 個の要素をソーティングするとき、最悪の場合の値の比較は合計何回必要か。

問 5 . 問 4 のソーティング方法は「マージソート (併合ソート)」と呼ばれている。より実用的なソーティングとして「クイックソート」が知られている。「クイックソート」のアルゴリズムを「マージソート」と比較して簡潔に説明し、これら 2 つのソーティング手法の利点・欠点を述べよ。

問題5：実験1～3に関する以下の問に答えよ。

レイソル博士は、柏キャンパスで新しい単細胞真核生物を単離して、それを *Kashiwanomyces inakae* (以下 *K. inakae*) と命名し、以下の性質を明らかにした。

K. inakae の一倍体には x 型と y 型の 2 つのタイプ (接合型) がある。両者を交配すると、融合して二倍体細胞を作った後に、減数分裂により 4 つの一倍体細胞の子孫を作る。減数分裂時には他の生物と同様に組み換えが起こる。

x 型と y 型の細胞が融合する際には、y 型細胞が分泌するフェロモン W 因子によって x 型細胞で遺伝子 *F* の転写が誘導される。

これらの知見を得たレイソル博士は、融合に関わるシグナル伝達経路を遺伝学的に調べることが目的に、遺伝子 *F* の発現誘導を指標に変異体を単離することを計画した。まず遺伝子 *F* のプロモーターによって発現が制御される大腸菌 *lacZ* (-ガラクトシダーゼ) 遺伝子を持つプラスミドを作成した。次に、変異剤処理を施した x 型細胞集団にこのプラスミドを導入して形質転換した。最後に、得られた形質転換体集団について、フェロモン W 因子存在下での -ガラクトシダーゼ活性を調べた。フェロモンによって *lacZ* が誘導されるものを *lacZ*(+)、されないものを *lacZ*(-)、フェロモンと無関係に恒常的に誘導されているものを *lacZ*(C) と表記することにした。

【実験1】

lacZ(-) の表現型を示す変異株 # 1 (x 型) では、突然変異によってこのシグナル伝達経路のどこかが障害されたと考えられる。しかし、それが 1ヶ所なのか複数なのかは分からない。これを探るために変異株 # 1 と y 型の野生株細胞の交配によって子孫の一倍体細胞集団 P を得た。

問1 - 1 . 変異株 # 1 が単一遺伝子 *G1* の変異による場合、集団 P 中で *lacZ*(-) の表現型を示す細胞の出現頻度は何%になると予想されるか？

問1 - 2 . 変異株 # 1 が示す *lacZ*(-) の表現型が 2 つの遺伝子 *G1* と *G10* の変異によるものであった場合、集団 P 中での *lacZ*(-) の表現型を示す細胞の出現頻度はどうなると考えられるか？ 2 つの変異が共存して初めて表現型が出る場合と、それぞれ単独でも表現型が出る場合に分けて考え、更に *G1* と *G10* の連鎖の影響も考慮に入れて答えよ。

【実験2】

実験1の結果、変異株#1は単一遺伝子 $G1$ の変異によることが示された。同様の解析によって、 $lacZ(C)$ の表現型を示す変異株#2(接合型 x) は単一遺伝子 $G2$ の変異によることが示された。この解析の過程で $G2$ に株#2と同じ変異を持つ y 型細胞株#2 y も得られた。変異株#1と#2の原因遺伝子である $G1$ と $G2$ のどちらが上流で機能するかを決める為に、変異株#1と#2 y との交配で一俵体細胞集団 Q を得た。なお $G1$ と $G2$ の間に連鎖はないことも分かっている。

問2 - 1 . $G1$ が $G2$ の上流で働く場合と下流で働く場合とに分けて、 $lacZ(+)$ 、 $lacZ(-)$ 、 $lacZ(C)$ それぞれの表現型を示す細胞の集団 Q における出現比率を予想し、その理由を述べよ。

【実験3】

変異体#3は遺伝子 $G3$ の変異によって $lacZ(C)$ の表現型を示している。この細胞の抽出液を SDS-ポリアクリルアミド電気泳動後、フィルターに写し取り、タンパク質 R に対する特異抗体を用いてウェスタンブロット解析を行った。その結果、タンパク質 R のバンドの位置が変異体#3では野生株に較べて高分子側にややシフトしており、その位置はフェロモン処理をした野生株のそれと等しいことが分かった。しかし細胞の抽出液を電気泳動前にアルカリフォスファターゼで処理をすると、このシフトは観察されなくなった。変異体#3からタンパク質 R を精製して質量分析を行ったところ、10番目の Ser 残基(Ser-10)にのみ翻訳後修飾が起こっていることが明らかになった。そこでレイソル博士は、この修飾のシグナル伝達における役割を探る為に、Ser-10をAlaに変化させる $R(S10A)$ 変異を持つ株を作成した。

問3 - 1 . タンパク質 R の Ser-10 にはどのような翻訳後修飾が起こっていると考えられるか？

問3 - 2 . Ala と Ser、そして上記の修飾を受けた Ser の構造式を書け。

問3 - 3 . $G3$ の恒常活性型変異と $R(S10A)$ 変異を二重に持つ x 型細胞の表現型は $lacZ(+)$ 、 $lacZ(-)$ 、 $lacZ(C)$ のいずれになると予想されるか、その理由とともに述べよ。

問3 - 4 . $G3$ が欠失した x 型細胞で遺伝子 F を恒常的に発現させるには、タンパク質 R の Ser-10 をどのようなアミノ酸に変えるとよいと思われるか、その理由とともに述べよ。

問題 6 : 20 塩基の 2 つの塩基配列 X, Y において、4 つの塩基は各々等確率で出現すると仮定する。 X と Y の間で、左から数えて同じ位置にある塩基をすべて比較するとき、以下の確率を各々求めよ。

1. X, Y が完全一致する確率。
2. X と Y の間で、1 つの塩基は一致せず、それ以外は全て一致し、かつ 15 塩基以上連続して一致する部分が存在する確率。
3. X と Y の間で、2 つの塩基は一致せず、それ以外は全て一致し、かつ 15 塩基以上連続して一致する部分が存在する確率。

問題 7 : 以下の文章における空欄を満たす数式および数値を答えよ。

細胞内における mRNA の分布が以下の 3 パターンに限定されると仮定する。

細胞内に 5 コピーある mRNA	5000 種類
細胞内に 50 コピーある mRNA	4500 種類
細胞内に 500 コピーある mRNA	500 種類

細胞中の mRNA の総数 N_{total} は (¹) 本である。細胞からランダムに S 本の mRNA を抽出したとき、特定の種類の mRNA (コピー数 C) を全く拾わない確率は、式 (²) で与えられる。コピー数 C の mRNA が全部で N_C 種類ある場合、それらが S 中で観測される期待値は (³) 種類となる。ただし、それぞれの生起確率が p である n 個のランダム変数が生起する回数の期待値が np で与えられるとする。ここで、 $N_C = 5000, C = 5, S = 5000$ を代入してみよう。近似式 $(1+1/x)^n \approx 1+n/x$ を用いると、コピー数 5 の mRNA は約 (⁴) 種しか観測できないことがわかる。